What is claimed:

1. An apparatus for retrieving documents, comprising:

a document dividing part configured to divide each document into character strings as index keys;

an index table configured to maintain the index keys and document information relating to each index key;

a query character string dividing part configured to divide a query character string into a plurality of index keys;

a retrieval condition analyzing part configured to analyze a retrieval condition including the index keys divided from the query character string and to generate a retrieval condition tree where the index keys are synthesized by at least one operator that retrieves an intermediate retrieval result including the document information from said index table; and

a retrieval condition evaluating part configured to evaluate each intermediate retrieval result obtained by the retrieval condition tree and to determine a final retrieval result, wherein:

said document dividing part divides the document into index keys of n-character strings having n characters and m-character strings having m characters where n is an integer greater than one and m is an integer less than n, and each of the m-character strings includes a last character of the document,

when at least two index keys are divided from the query character string by said query character string dividing part, said retrieval condition analyzing part includes:

a first condition tree generating part generating a first condition tree synthesized by at least one AND set operator obtaining an AND set of a plurality of the intermediate retrieval results based on said at least two index keys, and

a second condition tree generating part selecting a minimum number of index keys, which cover a full length of the query character string, from said at least two index keys and generating a second condition tree synthesized by at least one distance operator indicating a distance between appearance positions of said at least two index keys,

said retrieval condition analyzing part includes a document determining part obtaining candidate documents by executing the first condition tree and determining documents from the candidate documents by calculating a second condition tree, and

said first condition tree generating part generates the first condition tree by index keys used in the second condition tree and other index keys positioned in the query character string before or after the index keys used in the second condition tree and indicating a least number of the documents including the other index keys.

2. The apparatus as claimed in claim 1, wherein:

said query character string dividing part divides a query character string into more than two index keys of n-character strings having n characters to overlap a query character when a length of the query character string is more than n+1 characters where n is an integer greater than one; and

said retrieval condition analyzing part synthesizes the index keys by at least one distance operator indicating a distance between the index keys divided by said query character string dividing part.

3. The apparatus as claimed in claim 1, wherein:

said query character string dividing part defines a query character string as an index key when the query character string is n characters in length where n is an integer greater than one; and

said retrieval condition analyzing part generates a final retrieval condition formed by the index key.

4. The apparatus as claimed in claim 1, wherein:

said query character string dividing part outputs index keys from said index table where a first part of each index key identically corresponds to that of the query character string when a length of the query character string is less than n characters where n is an integer greater than one; and

said retrieval condition analyzing part generates the retrieval condition tree where the index keys, which are output by said query character string dividing part, are synthesized by at least one OR set operator obtaining an OR set of a plurality of the intermediate retrieval results.

5. The apparatus as claimed in claim 1, wherein said document dividing part divides the document into index keys of k-character strings having k characters where k, n and N are integers, n is equal to or greater than two, k is not less than one and is not more than N ($1 \leqq k \leqq N$), and the k-character string has k characters.

6. The apparatus as claimed in claim 1, wherein said document dividing part divides the document into index keys of k-character strings having k characters and m-character strings having m characters where k, m, n and N are integers, n and N are equal to or more than two and n is less than N ($n<N$), k is not less than n and is not more than N ($n \leqq k \leqq N$), and m is less than n ($m<n$).

7. The apparatus as claimed in claim 1, wherein said query character string dividing part outputs index keys from said index table where a beginning part of each index key identically corresponds to that of the query character string when a length of the query character string is less than n characters where n is an integer greater than one.

8. The apparatus as claimed in claim 1, wherein said document dividing part divides the document into successive character strings and said query character string dividing part divides the document into successive character strings, wherein each successive character string is formed by a single character type, and divides each successive character string into index keys by a method defined based on the single character type.

9. The apparatus as claimed in claim 8, wherein said document dividing part divides each successive character string formed by a single character type into index keys of n-character strings having n characters and m-character strings having m characters where n is an integer greater than one and m is an integer less than n, and each of the m-character strings includes a first character or a last character of each successive character string.

10. The apparatus as claimed in claim 8, wherein said document dividing part further extracts two-character strings and said query character string dividing part extracts two-character strings wherein each of two-character strings is formed by two different character types included in the document and predetermined as a combination character string.

11. The apparatus as claimed in claim 10, wherein said query character string dividing part does not extract a first character of the two-character string when one of the two different character types forming the two-character string forms the first character only.

12. The apparatus as claimed in claim 10, wherein said query character string dividing part does not extract a last character of the two-character string when one of the two different character types forming the two-character string forms the last character only.

13. The apparatus as claimed in claim 8, wherein said query character string dividing part outputs index keys formed by a single character string where a beginning part of each index key identically corresponds to the query character string when the query character string is formed by the single character type and a length of the query character string is equal to or less than a minimum length n defined for extracting a character string formed by the single character type.

14. The apparatus as claimed in claim 1, wherein when at least two index keys are divided from the query character string by said query character string dividing part, said retrieval condition analyzing part includes:

a first condition tree generating part generating a first condition tree synthesized by at least one AND set operator obtaining an AND set of a plurality of the intermediate retrieval results based on said at least two index keys; and

a second condition tree generating part selecting index keys, which cover a full length of the query character string and indicate a least total number of the documents including the index keys, from said at least two index keys and generating a second condition tree synthesized by at least one distance operator indicating a distance between appearance positions of said at least two index keys,

wherein:

said retrieval condition analyzing part includes a document determining part obtaining candidate documents by executing the first condition tree and determining documents from the candidate documents by calculating a second condition tree, and

said first condition tree generating part generates the first condition tree by index keys used in the second condition tree and other index keys positioned in the query character string before or after the index keys used in the second condition tree and indicating a least number of the documents including the other index keys themselves.

15. The apparatus as claimed in claim 1, wherein when a child node of an OR set operator obtaining an OR set of a plurality of retrieval results includes another OR set operator in said retrieval condition and a number of children nodes in said another OR set operator as a child node of the OR set operator is less than a threshold, said retrieval condition analyzing part includes a leveling part defining a latter child node as a former child node and eliminating factors of the latter child node from the former child node.

16. The apparatus as claimed in claim 1, wherein when a child node of an AND set operator obtaining an AND set of a plurality of retrieval results includes an OR set operator in said retrieval condition and a number of children nodes in the OR set operator as a child node of the OR set operator is less than a threshold after said retrieval condition is converted to another functionally equal retrieval condition defined by an OR operator which a child node includes an AND operator, said retrieval condition analyzing part executes to convert said retrieval condition.

17. The apparatus as claimed in claim 1, wherein said retrieval condition analyzing part synthesizes said first condition tree as a child node by an AND set operator to generate a synthesized first condition tree, and

said retrieval condition evaluating part obtains candidate documents based on the synthesized first condition tree and determines a final retrieval result.

18. The apparatus as claimed in claim 1, wherein said retrieval condition analyzing part additionally provides, in said first condition tree indicated by an AND set operator, an index key node as a child node of said AND set operator.

19. The apparatus as claimed in claim 1, wherein said retrieval condition evaluating part checks, in a set difference operator obtaining a set difference between two retrieval results, a first retrieved document obtained by a first node that is potentially a candidate document for a second node and determines the first retrieved document not to be the candidate document in accordance with a result of checking.

20. The apparatus as claimed in claim 1, wherein said retrieval condition evaluating part obtains, in order to evaluate an AND set operator, candidate documents for each child node, checks whether or not the candidate documents are included in a result set obtained by the AND set operator, determines whether or not the candidate documents are documents corresponding to the child node based on the check result, and adds the documents corresponding to the child node to the AND set operator based on the determination result.

21. A method for retrieving documents comprising the steps of:

(a) dividing each document into character strings as index keys;

(b) maintaining the index keys and document information relating to each index key;

(c) dividing a query character string into a plurality of index keys;

(d) analyzing a retrieval condition including the index keys divided from the query character string and generating a retrieval condition tree where the index keys are synthesized by at least one operator that retrieves an intermediate retrieval result including the document information from said index table; and

(e) evaluating each intermediate retrieval result obtained by the retrieval condition tree and determining a final retrieval result, wherein:

said step (a) divides the document into index keys of n-character strings having n characters and m-character strings having m characters where n is an integer greater than one and m is an integer less than n, and each of the m-character strings includes a last character of the document,

when at least two index keys are divided from the query character string in said step (c), said step (d) includes the steps of:

(f) generating a first condition tree synthesized by at least one AND set operator obtaining an AND set of a plurality of the intermediate retrieval results based on said at least two index keys, and

(g) selecting a minimum number of index keys, which cover a full length of the query character string, from said at least two index keys and generating a second condition tree synthesized by at least one distance operator indicating a distance between appearance positions of said at least two index keys,

said step (d) includes a step of obtaining candidate documents by executing the first condition tree and determining documents from the candidate documents by calculating a second condition tree, and

said step (f) generates the first condition tree by index keys used in the second condition tree and other index keys positioned in the query character string before or after the index keys used in the second condition tree and indicating a least number of the documents including the other index keys.

22. The method as claimed in claim 21, wherein:

said step (c) divides a query character string into more than two index keys of n-character strings having n characters to overlap query character when a length of the query character string is more than n+1 characters where n is an integer greater than one, and

said step (d) synthesizes the index keys by at least one distance operator indicating a distance between the index keys divided in said step (c).

23. The method as claimed in claim 21, wherein:

said step (c) defines a query character string as an index key when the query character string is n characters in length where n is an integer greater than one; and

said step (d) generates a final retrieval condition formed by the index key.

24. The method as claimed in claim 21, wherein:

said step (c) outputs index keys from said index table where a first part of each index key identically corresponds to that of the query character string when a length of the query character string is less than n characters where n is an integer greater than one; and

said step (d) generates the retrieval condition tree where the index keys, which are output in said step (c), are synthesized by at least one OR set operator obtaining an OR set of a plurality of the intermediate retrieval results.

25. The method as claimed in claim 21, wherein said step (d) synthesizes said first condition tree as a child node by an AND set operator to generate a synthesized first condition tree, and said step (e) obtains candidate documents based on the synthesized first condition tree and determines a final retrieval result.

26. The method as claimed in claim 25, wherein said step (d) additionally provides, in said first condition tree indicated by an AND set operator, an index key node as a child node of said AND set operator.

27. The method as claimed in claim 21, wherein said step (e) checks, in a set difference operator obtaining a set difference between two retrieval results, a first retrieved document obtained by a first node that is potentially a candidate document for a second node and determines the first retrieved document not to be the candidate document in accordance with a result of checking.

28. The method as claimed in claim 21, wherein said step (e) obtains, in order to evaluate an AND set operator, candidate documents for each child node, checks whether or not the candidate documents are included in a result set obtained by the AND set operator, determines whether or not the candidate documents are documents corresponding to the child node based on the check result, and adds the documents corresponding to the child node to the AND set operator based on the determination result.

29. A computer-readable recording medium having a program code recorded therein for causing a computer to retrieve documents, said program code comprising codes for:

(a) dividing each document into character strings as index keys;

(b) maintaining the index keys and document information relating to each index key;

(c) dividing a query character string into a plurality of index keys;

(d) analyzing a retrieval condition including the index keys divided from the query character string and generating a retrieval condition tree where the index keys are synthesized by at least one operator that retrieves an intermediate retrieval result including the document information from said index table; and

(e) evaluating each intermediate retrieval result obtained by the retrieval condition tree and determining a final retrieval result, wherein:

said code (a) divides the document into index keys of n-character strings having n characters and m-character strings having m characters where n is an integer greater than one and m is an integer less than n, and each of m-character strings includes a last character of the document,

when at least two index keys are divided from the query character string by said code (c), said code (d) includes codes of:

(f) generating a first condition tree synthesized by at least one AND set operator obtaining an AND set of a plurality of the intermediate retrieval results based on said at least two index keys,

(g) selecting a minimum number of index keys, which cover a full length of the query character string, from said at least two index keys and generating a second condition tree synthesized by at least one distance operator indicating a distance between appearance positions of said at least two index keys,

said code (d) includes a code of obtaining candidate documents by executing the first condition tree and determining documents from the candidate documents by calculating a second condition tree, and

said code (f) generates the first condition tree by index keys used in the second condition tree and other index keys positioned in the query character string before or after the index keys in the second condition tree and indicating a least number of the documents including the other index keys.

30. The computer-readable recording medium as claimed in claim 29, wherein said code (c) divides a query character string into more than two index keys of n-character strings having n characters to overlap a query character when a length of the query character string is more than n+l characters where n is an integer greater than one, and

said code (d) synthesizes the index keys by at least one distance operator indicating a distance between the index keys divided by said code (c).

31. The computer-readable recording medium as claimed in claim 29, wherein said code (c) defines a query character string as an index key when the query character string is n characters in length where n is an integer greater than one, and

said code (d) generates a final retrieval condition formed by the index key.

32. The computer-readable recording medium as claimed in claim 29, wherein said code (c) outputs index keys from said index table where a first part of each index key identically corresponds to that of the query character string when a length of the query character string is less than n characters where n is an integer greater than one, and

said code (d) generates the retrieval condition tree where the index keys, which are output by said code (c), are synthesized by at least one OR set operator obtaining an OR set of a plurality of the intermediate retrieval results.

33. The computer-readable recording medium as claimed in claim 29, wherein said code (d) synthesizes said first condition tree as a child node by an AND set operator to generate a synthesized first condition tree, and

said code (e) obtains candidate documents based on the synthesized first condition tree and determines a final retrieval result.

34. The computer-readable recording medium as claimed in claim 33, wherein said code (d) additionally provides, in said first condition tree indicated by an AND set operator, an index key node as a child node of said AND set operator.

35. The computer-readable recording medium as claimed in claim 29, wherein said code (e) checks, in a set difference operator obtaining a set difference between two retrieval results, a first retrieved document obtained by a first node that is potentially a candidate document for a second node and determines the first retrieved document not to be the candidate document in accordance with a result of checking.

36. The computer-readable recording medium as claimed in claim 29, wherein said code (e) obtains, in order to evaluate an AND set operator, candidate documents for each child node, checks whether or not the candidate documents are included in a result set obtained by the AND set operator, determines whether or not the candidate documents are documents corresponding to the child node based on the check result, and adds the documents corresponding to the child node to the AND set operator based on the determination result.

37. The apparatus as claimed in claim 1, wherein said retrieval condition analyzing part synthesizes said first condition tree as a child node by an AND set operator to generate a synthesized first condition tree, and

said retrieval condition evaluating part obtains candidate documents based on the synthesized first condition tree and determines a final retrieval result.

38. The apparatus as claimed in claim 1, wherein said retrieval condition evaluating part checks, in a set difference operator obtaining a set difference between two retrieval results, a first retrieved document obtained by a first node possible to be a candidate document for a second node and determines the first retrieved document not to be the candidate document in accordance with a result of checking.

39. The apparatus as claimed in claim 14, wherein said retrieval condition evaluating part checks, in a set difference operator obtaining a set difference between two retrieval results, a first retrieved document obtained by a first node possible to be a candidate document for a second node and determines the first retrieved document not to be the candidate document in accordance with a result of checking.